Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study

Leijenaar, Ralph T H ; Bogowicz, Marta ; Jochems, Arthur ; Hoebers, Frank J P ; Wesseling, Frederik W R ; Huang, Sophie H ; Chan, Biu ; Waldron, John N ; O'Sullivan, Brian ; Rietveld, Derek ; Leemans, C Rene ; Brakenhoff, Ruud H ; Riesterer, Oliver ; Tanadini-Lang, Stephanie ; Guckenberger, Matthias ; Ikenberg, Kristian ; Lambin, Philippe

Abstract: OBJECTIVES Human papillomavirus (HPV) positive oropharyngeal cancer (oropharyngeal squamous cell carcinoma, OPSCC) is biologically and clinically different from HPV negative OPSCC. Here, we evaluate the use of a radiomic approach to identify the HPV status of OPSCC. METHODS Four independent cohorts, totaling 778 OPSCC patients with HPV determined by p16 were collected. We randomly assigned 80% of all data for model training (N = 628) and 20% for validation (N = 150). On the pre-treatment CT images, 902 radiomic features were calculated from the gross tumor volume. Multivariable modeling was performed using least absolute shrinkage and selection operator. To assess the impact of CT artifacts in predicting HPV (p16), a model was developed on all training data (M) and on the artifact-free subset of training data (M). Models were validated on all validation data (V), and the subgroups with (V) and without (V) artifacts. Kaplan-Meier survival analysis was performed to compare HPV status based on p16 and radiomic model predictions. RESULTS The area under the receiver operator curve for M and M ranged between 0.70 and 0.80 and was not significantly different for all validation data sets. There was a consistent and significant split between survival curves with HPV status determined by p16 [p = 0.007; hazard ratio (HR): 0.46], M (p = 0.036; HR: 0.55) and M (p = 0.027; HR: 0.49). CONCLUSION This study provides proof of concept that molecular information can be derived from standard medical images and shows potential for radiomics as imaging biomarker of HPV status. Advances in knowledge: Radiomics has the potential to identify clinically relevant molecular phenotypes.

DOI: https://doi.org/10.1259/bjr.20170498

DOI: https://doi.org/10.1259/bjr.20170498

BJR https://doi.org/10.1259/bjr.20170498

Received: 03 July 2017 | Revised: 17 January 2018 | Accepted: 12 February 2018

Cite this article as:
Leijenaar RTH, Bogowicz M, Jochems A, Hoebers FJP, Wesseling FWR, Huang SH, et al. Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study. *Br J Radiol* 2018; **91**: 2017049811075.

FULL PAPER

Development and validation of a radiomic signature to predict HPV (p16) status from standard CT imaging: a multicenter study

[1]RALPH TH LEIJENAAR, PhD, [2]MARTA BOGOWICZ, MSc, [1]ARTHUR JOCHEMS, PhD, [3]FRANK JP HOEBERS, PhD, MD, [3]FREDERIK WR WESSELING, MSc, MD, [4]SOPHIE H HUANG, MSc, MRT(T), MD, [4]BIU CHAN, MSc, MRTT, [4]JOHN N WALDRON, MSc, MD, FRCPC, [4]BRIAN O'SULLIVAN, MBBCh, BAO, FRCPC, [5]DEREK RIETVELD, MSc, MD, [6]C RENE LEEMANS, PhD, MD, [6]RUUD H BRAKENHOFF, PhD, MD, [2]OLIVER RIESTERER, MD, PD, [2]STEPHANIE TANADINI-LANG, PhD, [2]MATTHIAS GUCKENBERGER, MD, [7]KRISTIAN IKENBERG, MD and [1]PHILIPPE LAMBIN, PhD, MD

[1]The D-Lab: Decision Support for Precision Medicine GROW - School for Oncology and Developmental Biology & MCCC Maastricht University Medical Centre+ Maastricht, Maastricht, Netherlands
[2]Department of Radiation Oncology, University Hospital Zurich and University of Zurich, Zurich, Switzerland
[3]Department of Radiation Oncology, GROW- School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, Netherlands
[4]Department of Radiation Oncology, Princess Margaret Cancer Center, University of Toronto, Toronto, Ontario, Canada
[5]Department of Radiation Oncology, VU University Medical Center, Amsterdam, Netherlands
[6]Department of Otolaryngology/Head and Neck Surgery, VU University Medical Center, Amsterdam, Netherlands
[7]Institute of Pathology and Molecular Pathology, University Hospital Zurich and University of Zurich, Zurich, Switzerland

Address correspondence to: Dr Ralph TH Leijenaar
E-mail: *ralph.leijenaar@maastrichtuniversity.nl*

The authors Ralph TH Leijenaar and Marta Bogowicz contributed equally to the work.

Objectives: Human papillomavirus (HPV) positive oropharyngeal cancer (oropharyngeal squamous cell carcinoma, OPSCC) is biologically and clinically different from HPV negative OPSCC. Here, we evaluate the use of a radiomic approach to identify the HPV status of OPSCC.

Methods: Four independent cohorts, totaling 778 OPSCC patients with HPV determined by p16 were collected. We randomly assigned 80% of all data for model training (N = 628) and 20% for validation (N = 150). On the pre-treatment CT images, 902 radiomic features were calculated from the gross tumor volume. Multivariable modeling was performed using least absolute shrinkage and selection operator. To assess the impact of CT artifacts in predicting HPV (p16), a model was developed on all training data (M_{all}) and on the artifact-free subset of training data ($M_{no\ art}$). Models were validated on all validation data (V_{all}), and the subgroups with (V_{art}) and without ($V_{no\ art}$) artifacts. Kaplan–Meier survival analysis was performed to compare HPV status based on p16 and radiomic model predictions.

Results: The area under the receiver operator curve for M_{all} and $M_{no\ art}$ ranged between 0.70 and 0.80 and was not significantly different for all validation data sets. There was a consistent and significant split between survival curves with HPV status determined by p16 [p = 0.007; hazard ratio (HR): 0.46], M_{all} (p = 0.036; HR: 0.55) and $M_{no\ art}$ (p = 0.027; HR: 0.49).

Conclusion: This study provides proof of concept that molecular information can be derived from standard medical images and shows potential for radiomics as imaging biomarker of HPV status.

Advances in knowledge: Radiomics has the potential to identify clinically relevant molecular phenotypes.

INTRODUCTION

Over the last years, the incidence of oropharyngeal squamous cell carcinoma (OPSCC) has shown a dramatic increase relative to other head and neck cancers, with a substantial proportion of OPSCC being linked to human papillomavirus (HPV) infections.[1]

HPV positive OPSCC is biologically and clinically different from HPV negative OPSCC, which is often related to

alcohol and tobacco consumption. HPV positive OPSCC has been shown to have superior response to radiochemotherapy. Approximately, 80% of HPV positive OPSCC patients achieve locoregional control and 5 years overall survival, in comparison to less than 50% of patients with HPV negative OPSCC and non-oropharyngeal head and neck cancers.[2,3] This favorable outcome makes HPV positive OPSCC in particular interesting for de-escalation protocols.[4]

Widely accepted methods for detection of HPV infection are *in situ* hybridization for viral DNA, HPV DNA or RNA PCR, and immunohistochemical investigation of the level of p16 expression, which strongly correlates with HPV infection.[5]

Radiomics is a rapidly emerging field, introduced in 2012, which concerns with the high-throughput mining of large amounts of quantitative features, derived from (standard-of-care) medical imaging, for knowledge extraction.[6–9] Radiomics is in particular promising within decision support systems for precision medicine[10–12] and its potential to predict HPV status in head and neck cancer has been recognized.[13] Indeed, previous studies have reported radiologic differences between HPV positive and negative cases in terms of qualitative radiologist's readout[14] or perfusion CT.[15] Furthermore, exploratory radiomic studies have shown that heterogeneity of image-based density is potentially associated with HPV in OPSCC.[16,17]

Most of the studies investigating imaging phenotypes of tumors are based on single center data, which introduce bias to a model and limits its applicability.[8] In particular, factors such as CT scanner, tube voltage, tube current, reconstruction kernel and contrast agent influence the results of quantitative analysis. In this multicenter study, we further investigate if a quantitative CT-based radiomic approach can objectively identify the HPV (p16) status of OPSCC, by developing and validating a radiomic signature on a large and international collection of patient data from four different institutions. This study does not intend to develop methodology to replace existing HPV tests, yet aims to provide a proof of concept that radiomics is able to derive molecular information from standard medical images.

METHODS AND MATERIALS

Patients and CT imaging

Four independent cohorts, with a total of 778 OPSCC patients with HPV status determined by p16 immunohistochemistry and treated with curative intent by radiation therapy with/without concurrent chemotherapy, were collected from the Princess Margaret Cancer Center (N = 427), the VU University Medical Center (N = 158), the University Hospital Zürich (N = 100) and MAASTRO clinic (N = 93). All patients underwent pre-treatment contrast enhanced CT imaging of the head and neck. The gross primary tumor volume (GTV) was manually delineated for each patient for treatment planning purposes. The delineation was not standardized and was performed according to the clinical protocols, separate for each institute. Images were visually assessed for the presence of CT artifacts (*e.g.* streak artifacts due to dental fillings) within the GTV. A more detailed description of acquired CT images for each of the included cohorts can be found in the Supplementary Material 1. Institutional review board approval was obtained for each of the participating centers. Patients provided informed written consent, unless the need for written consent for this retrospective study was waived by the participating center.

Image analysis

Prior to analysis, all images were resampled to isotropic voxels of 2 mm, using linear interpolation.[18] A total of 902 radiomic features were calculated, divided into five groups: tumor intensity, shape, texture, Wavelet and Laplacian of Gaussian. All features were extracted using in-house developed software, using Matlab 2014a (MathWorks, Natick, MA). Feature descriptions and mathematical definitions can be found elsewhere.[8,19] To calculate wavelet features, we used the low pass approximation and the high pass decomposition (*i.e.* applying either a low or high pass filter in each direction, respectively), since these decompositions are directionally invariant. For Laplacian of Gaussian features, the texture size (fine to coarse) was highlighted by modifying the Gaussian radius parameter from 2 to 7 mm with 1 mm increments. Textural features were computed discretizing image intensities into bins, using both a bin width of 10 and 25 Hounsfield unit.[20]

Statistical analysis

We randomly assigned 80% of all data for model training (N = 628) and 20% for validation (N = 150), with balanced HPV status, institution, and number of patients with visible CT artifacts.

Highly correlated features were first removed from further analysis by evaluating all pairwise correlations in the training data set. For each highly correlated feature pair (Pearson correlation coefficient $\rho > 0.9$), the variable with the largest mean absolute correlation with all remaining features was removed.

Multivariable logistic regression was performed using the least absolute shrinkage and selection operator model selection technique,[21] with 100 times repeated 10-fold cross-validation to select the optimal tuning parameter (λ). To further reduce the chance of overfitting on the training data, we selected the simplest candidate model (*i.e.* the model with fewest non-zero coefficients), *i.e.* within one standard error of the best performing model. The area (AUC) under the receiver operator curve (ROC) was used to assess model performance in predicting HPV (p16) status.

Finally, we compared Kaplan–Meier survival curves between patients with positive and negative HPV status, based on conventional p16 immunohistochemistry and based on radiomic model HPV predictions, for all validation patients. Model class predictions were made with a probability cutoff of 0.5. Overall survival was defined as the time from start of treatment to death as a result of any cause. A log-rank test was applied to test for significant differences between survival curves.

Table 1. HPV status, presence of CT artifacts and median follow-up time for the PMH, the VUmc, the USZ, MAASTRO clinic (MAASTRO), the 80% training data (training) and the 20% validation data (validation)

Variable	PMH (n = 427)	VUmc (n = 158)	USZ (n = 100)	MAASTRO (n = 93)	Training (n = 628)	Validation (n = 150)
HPV (p16) status						
Positive	303 (71%)	34 (22%)	56 (56%)	33 (35%)	344 (55%)	82 (55%)
Negative	124 (29%)	124 (78%)	44 (44%)	60 (65%)	284 (45%)	68 (45%)
CT artifacts						
Yes	219 (51%)	69 (44%)	57 (57%)	26 (28%)	300 (48%)	71 (47%)
No	208 (49%)	89 (56%)	43 (43%)	67 (72%)	328 (52%)	79 (53%)
Overall survival						
Median follow-up (months)	71.6	74	44.5	51.8	69.4	65.1
Cohort						
PMH	–	–	–	–	343 (55%)	84 (56%)
VUmc	–	–	–	–	128 (20%)	30 (20%)
USZ	–	–	–	–	82 (13%)	18 (12%)
MAASTRO	–	–	–	–	75 (12%)	18 (12%)

PMH, Princess Margaret Cancer Center; UHZ, University Hospital Zürich; VUmc, VU University Medical Center.
For the training and validation data sets, the amount of patients from each individual cohort are given as well. Median follow-up for overall survival was determined by "reverse" Kaplan–Meier analysis (*i.e.* inversed censoring).

To assess the impact of CT artifacts, a model was also developed on the subset of patients in the training cohort for which there were no visible CT artifacts within the GTV. All model validation was subsequently performed on the entire validation data (V_{all}), and the subgroups of validation patients with (V_{art}) and without ($V_{no\ art}$) CT artifacts. AUC values for paired ROC curves were compared using DeLong's test.[22] Model calibration was measured by the intercept and slope of the logistic calibration curve.[23]

To further compare the two models, confusion matrices for HPV (p16) predictions by M_{all} and $M_{no\ art}$ on all validation data V_{all} were determined.

Statistical analysis was performed in R (R Foundation for Statistical Computing;v. 3.3.3).

RESULTS

Radiomic models for HPV prediction

Patient characteristics, including HPV status, presence of CT artifacts, and follow-up time are summarized in Table 1.

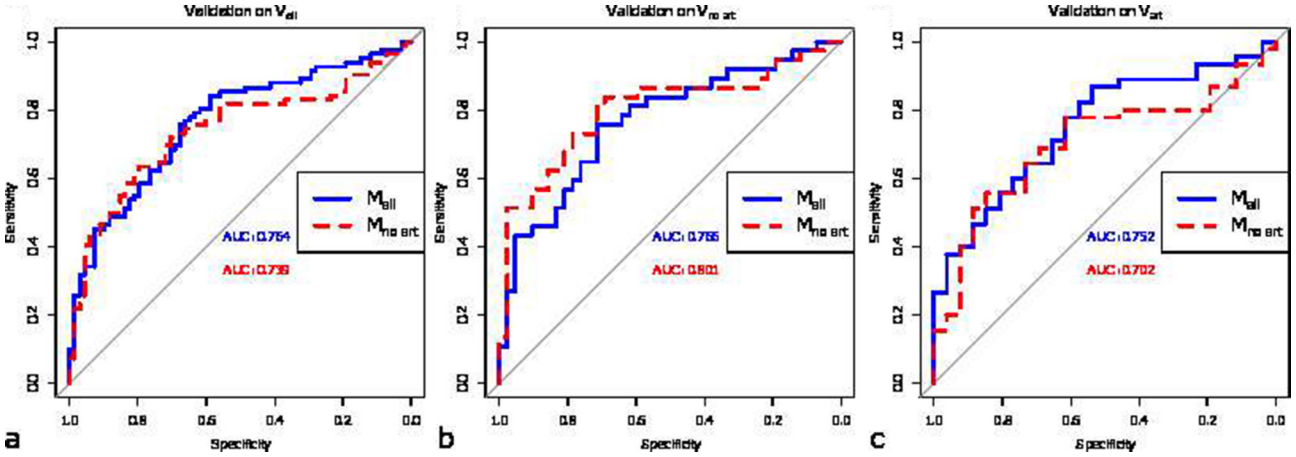
A subset of 165 uncorrelated features was preselected for the full training cohort, whereas a subset of 173 uncorrelated features was identified for the training data without CT artifacts. The models developed on all training data (M_{all}; 37 degrees of freedom) and on the subset of training data without CT artifacts ($M_{no\ art}$; 50 degrees of freedom). For model performance on the training data, we observed an AUC of 0.824 [95% CI (0.791–0.856)]. and 0.868 [95% CI (0.830–0.906)] for M_{all} and $M_{no\ art}$, respectively. Both models were subsequently validated on V_{all}, $V_{no\ art}$ and V_{art}. The resulting AUC values, logistic calibration intercepts and

Table 2. AUC values, logistic calibration intercepts and slopes for the model developed on all training data (M_{all}) and the model developed on the subset of training patients without CT artifacts ($M_{no\ art}$), validated in all validation data (V_{all}), the subset of validation data without CT artifacts ($V_{no\ art}$) and the subset of validation data with CT artifacts (V_{art})

Model	Validation dataset	AUC	Intercept	Slope
M_{all}	V_{all}	0.7636 [95% CI (0.6874–0.8399)]	0.034	1.041
	$V_{no\ art}$	0.7658 [95% CI 0.6592–0.8724)]	–0.238	1.191
	V_{art}	0.7521 [95% CI (0.6378–0.8665)]	0.37	0.852
$M_{no\ art}$	V_{all}	0.7391 [95% CI (0.6582–0.8199)]	0.408	0.561
	$V_{no\ art}$	0.8005 [95% CI (0.6967–0.9044)]	0.057	1.103
	V_{art}	0.7017 [95% CI (0.5775–0.8259)]	0.767	0.341

AUC, area under the curve.

Figure 1. ROC plots for the model developed on all training data M_{all} and the model developed on the subset of training data without CT artifacts $M_{no\ art}$, validated on all validation data V_{all} (a), the subset of validation data without CT artifacts $V_{no\ art}$ (b) and the subset of validation data with CT artifacts V_{art} (c). AUC, area under the curve; ROC, receiver operator curve.



slopes are summarized in Table 2. The corresponding ROC plots are shown in Figure 1.

Kaplan–Meier survival curves, including numbers at risk, for all validation data V_{all} are shown in Figure 2. For HPV determined by p16, there was a significant split between survival curves for HPV (p16) positive and negative cases ($p = 0.007$), with a hazard ratio of 0.46 [95% CI (0.26–0.82)]. For HPV (p16) predictions by M_{all} ($p = 0.036$) and $M_{no\ art}$ ($p = 0.027$), we observed a similar significant split between survival curves, with hazard ratios of 0.55 [95% CI (0.31–0.97)] and 0.49 [95% CI (0.26–0.93)], respectively.

Comparison of HPV models based on training data with and without CT artifacts

AUC values for HPV (p16) predictions made by M_{all} and $M_{no\ art}$ were not significantly different for all validation data sets. Confusion matrices for HPV (p16) predictions by M_{all} and $M_{no\ art}$ on all validation data V_{all} are shown in Tables 3 and 4.

DISCUSSION

In this multicenter study, we developed and validated a CT based radiomic signature to predict the HPV status of OPSCC patients. In the context of radiogenomics,[24,25] our study provides a proof of concept that molecular information can be inferred from standard medical images by means of radiomics.

Previous exploratory radiomic studies that indicated a correlation between HPV infection and heterogeneity of imaging-based tumor density in OPSCC[16,17] either were performed on small populations without validation, or only used single institution data for both model development and validation. This is a major issue in radiomic studies, as can be learned from recent literature describing the process and challenges of radiomics.[8,13,26–28]

Figure 2. Kaplan–Meier curves and number of patients at risk for HPV predictions by M_{all} *vs* p16 (a) and $M_{no\ art}$ *vs* p16 (b). Survival times are in months. HPV, human papilloma virus.

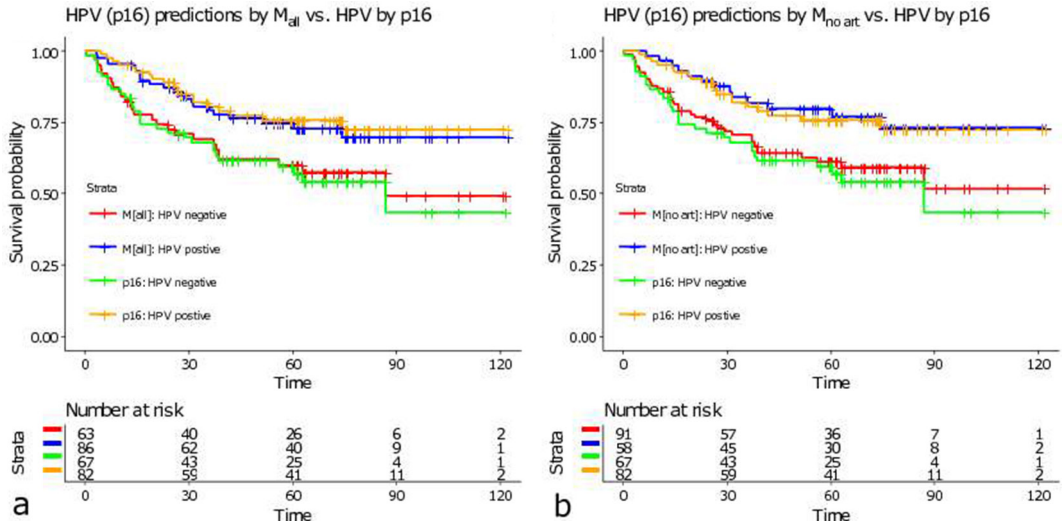


Table 3. Confusion matrix for HPV (p16) predictions by M_{all} on all validation data V_{all}

		Reference	
		HPV–	HPV+
Predictions	HPV–	44	19
	HPV+	24	63

HPV, human papillomavirus.

In this multicenter study, we used a large collection of imaging data from four different institutions for model development and validation.

Previously published studies report that HPV positive tumors are more homogenous in CT density.[16,17] Although a full interpretation of the complex radiomic signature for HPV is difficult, we have analyzed the meaning of several features selected both in the model trained on all training data and the model trained on the subset of training data with no CT artifacts. We have summarized our observations in Table 5. Our models selected different features than the ones previously published. However, as many radiomic features are correlated with each other, it is more relevant to compare features interpretation than name. Our multicenter results confirm that HPV tumors are more homogenous in CT density. This study provides also an additional insight into HPV imaging phenotype. We have observed that the HPV positive tumors seem to be characterized by lower contrast uptake, lower minimum density, and higher changes in the intensity of adjacent voxels.

Histopathology analysis shows differences between HPV+ and HPV- microscopic traits. For example, HPV+ tumors are characterized by lobular growth, infiltrating lymphocytes and well differentiated cells.[29] A direct link between these histopathology traits and radiomic signature of HPV positive tumors is not possible at this stage, and would require further investigation on a surgical cohort with full histology data.

Including data from different institutions introduces variety in image acquisition and reconstruction, which has been shown to affect radiomic features.[30,31] Shafiq-ul-Hassan et al[18] investigated voxel-size dependency of radiomic features and found that the robustness of radiomic analyses can be improved by resampling to a nominal voxel size or by normalizing the voxel size. All images in this study were, therefore, resampled to isotropic voxels of 2 mm, which was approximately the average slice spacing, using linear interpolation. Furthermore, as shown previously, textural features and their interpretation are affected by the bin width used to discretize image intensities[20] . Therefore, features calculated for different bin widths may provide additional predictive information. To account for this, textural features were computed using both a bin width of 10 and 25 Hounsfield unit. In our HPV radiomic signatures, features calculated using both bin sizes were selected and no preference for a bin size was observed.

Table 4. Confusion matrix for HPV (p16) predictions by $M_{no\ art}$ on all validation data V_{all}

		Reference	
		HPV–	HPV+
Predictions	HPV–	57	34
	HPV+	11	48

HPV, human papillomavirus.

Besides variability in CT imaging, demographic differences also have to be considered. Developing a model on a single, independent cohort is, therefore, unlikely to sufficiently capture the variability that exists across datasets, resulting in a model with poor generalizability. We, therefore, performed our model development on more heterogeneous data, by randomly assigning 80% of all included data for model training and 20% for testing, with balanced HPV (p16) status, institution and number of patients with visible CT artifacts.

A common concern in the analysis of CT images of head and neck cancer are metallic dental fillings or other high atomic number material implants, which result in imaging artifacts.[32] An existing radiomic signature for overall survival[7] has previously been shown to have prognostic power regardless of CT artifacts.[33] Another recent study exploring the link between HPV status and CT radiomics, preprocessed images by completely removing artifacts affected slices from analysis.[17] However, such a process neglects potentially relevant three-dimensional information. To investigate the impact of CT artifacts on HPV prediction, we developed a model on all data (M_{all}) and a model on the subset of data without artifacts ($M_{no\ art}$). What can be observed from our results is that there is no significant difference in discriminative power of both models. However, overall calibration of M_{all} was better than that of $M_{no\ art}$. It has to be noted that the extent of CT artifacts and the impact on radiomic features will vary between patients. For an individual patient, model accuracy will, therefore most likely depend on the amount of the tumor region that is obscured by artifacts. This would have to be further investigated, preferably including techniques for metal artifact reduction in CT.

Since HPV-related OPSCC have been shown to have superior response to radio-chemotherapy,[2,3] we compared Kaplan–Meier survival curves between patients with positive and negative HPV status, based on p16 and model class predictions by M_{all} and $M_{no\ art}$ for all validation patients. Indeed, we observed a significant split ($p < 0.05$) between HPV positive and HPV negative patients based on p16. For HPV (p16) predictions based on both models, we obtained survival curves similar to that of p16, with significantly different survival for HPV positive and HPV negative patients, indicating that model predictions are indeed in line with p16.

It has previously been shown that part of the OPSCC patients that test positive for p16 immunohistochemistry are in fact HPV DNA negative.[34,35] Since HPV testing for patients included in our study was performed by p16, the likelihood of false-positives has to be acknowledged. Furthermore, model class predictions

Table 5. Interpretation of selected features in the HPV radiomic signatures

	Radiomic feature	Interpretation
Decreased in the HPV tumors in comparison to HPV negative	Gray level size zone matrix Small zone emphasis	Higher homogeneity
	Gray level co-occurrence matrix Inverse variance	Higher changes in the intensity of adjacent voxels
	Laplacian of Gaussian (4 mm) 10^{th} percentile	Lower minimum HU value
Increased in the HPV tumors in comparison to HPV negative	Gray level size zone matrix Low gray level large size emphasis	Lower contrast uptake
	Laplacian of Gaussian (3 mm) kurtosis	More outliers

HPV, human papillomavirus; HU, Hounsfield unit.

(*i.e.* predicting either HPV positive or HPV negative), were made with a probability cut-off of 0.5, meaning that the costs for false-positives and false-negatives were considered equal. In clinical practice, this will not be the case and false-positives should be avoided (*i.e.* have a high cost), in order not to unjustly deescalate any patient's treatment. To achieve a clinically acceptable level of accuracy, further development and validation would be needed, including HPV DNA. It is, therefore, important to note that radiomic HPV prediction models are not meant to replace HPV testing and we acknowledge that clinical decision making should always be made on the universally accepted most accurate testing (*i.e.* p16 positive tumors should be subjected to HPV DNA testing).

Considering HPV testing is routinely performed for OPSCC patients in western countries, the clinical usefulness of a radiomic biomarker could be considered to be limited. However, our results show there is potential for radiomics to serve as a cost-effective, complementary method for HPV screening, which may also be useful in non-oropharyngeal SCCs.[36] Another potential application for a reliable radiomic biomarker could be to perform (retrospective) HPV analyses when no tissue samples are available, or in countries where it is not routinely done.

In this study, we only considered the primary tumor. However, HPV-associated OPSCC, commonly present with a relatively smaller primary tumor, and relatively more advanced nodal disease. Severity of the disease may then be overestimated by the resulting higher tumour, node and metastasis and overall stage, as these are related to other head and neck SCC.[37] HPV has also been shown to affect the morphology of affected lymph nodes.[38] Including radiomics of involved lymph nodes could potentially provide additional value in predicting HPV status. Furthermore, additional improvement in inferring tumor HPV status may be achieved when combining radiomics with clinical features.[14]

DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Ralph Leijenaar is CTO of OncoRadiomics SA. Ralph Leijenaar and Philippe Lambin are co-inventors of patents related to radiomics.

FUNDING

Authors acknowledge financial support from the ERC advanced grant (ERC-ADG-2015, n° 694812 - Hypoximmuno), the QuIC-ConCePT project (IMI JU; grant no. 115151). This research is also supported by the Dutch technology Foundation STW (grant n° 10696 DuCAT & n° P14-19 Radiomics STRaTegy), which is the applied science division of NWO, and the Technology Programme of the Ministry of Economic Affairs. Authors also acknowledge financial support from the EU 7th framework program (ARTFORCE - n° 257144, REQUITE - n° 601826), SME Phase 2 (EU proposal 673780 – RAIL), EUROSTARS (DART, EDECIDE), the European Program H2020-2015-17 (BD2Decide - PHC30-689715 and ImmunoSABR - n° 733008), Interreg V-A Euregio Meuse-Rhine ("Euradiomics"), Alpe d'HuZes-KWF (DESIGN), Kankeronderzoekfonds Limburg from the Health Foundation Limburg, the Zuyderland-MAASTRO grant and the Dutch Cancer Society.

REFERENCES

1. Pytynia KB, Dahlstrom KR, Sturgis EM. Epidemiology of HPV-associated oropharyngeal cancer. *Oral Oncol* 2014; **50**: 380–6. doi: https://doi.org/10.1016/j.oraloncology.2013.12.019
2. Ang KK, Harris J, Wheeler R, Weber R, Rosenthal DI, Nguyen-Tân PF, et al. Human papillomavirus and survival of patients with oropharyngeal cancer. *N Engl J Med* 2010; **363**: 24–35. doi: https://doi.org/10.1056/NEJMoa0912217
3. Lassen P, Primdahl H, Johansen J, Kristensen CA, Andersen E, Andersen LJ, et al. Impact of HPV-associated p16-expression on radiotherapy outcome in advanced oropharynx and non-oropharynx cancer. *Radiother Oncol* 2014; **113**: 310–6. doi: https://doi.org/10.1016/j.radonc.2014.11.032
4. Rietbergen MM, Brakenhoff RH, Bloemena E, Witte BI, Snijders PJ, Heideman DA, et al. Human papillomavirus detection and comorbidity: critical issues in selection of patients with oropharyngeal cancer for treatment De-escalation trials. *Ann Oncol* 2013;

24: 2740–5. doi: https://doi.org/10.1093/annonc/mdt319
5. Krupar R, Hartl M, Wirsching K, Dietmaier W, Strutz J, Hofstaedter F. Comparison of HPV prevalence in HNSCC patients with regard to regional and socioeconomic factors. *Eur Arch Otorhinolaryngol* 2014; **271**: 1737–45. doi: https://doi.org/10.1007/s00405-013-2693-8
6. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer* 2012; **48**: 441–6. doi: https://doi.org/10.1016/j.ejca.2011.11.036
7. Aerts HJ, Velazquez ER, Leijenaar RT, Parmar C, Grossmann P, Carvalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat Commun* 2014; **5**: 4006. doi: https://doi.org/10.1038/ncomms5006
8. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 2017; **14**: 749–62. doi: https://doi.org/10.1038/nrclinonc.2017.141
9. Radiomics. 2018. Available from: http://www.radiomics.world/.
10. Lambin P, van Stiphout RG, Starmans MH, Rios-Velazquez E, Nalbantov G, Aerts HJ, et al. Predicting outcomes in radiation oncology-multifactorial decision support systems. *Nat Rev Clin Oncol* 2013; **10**: 27–40. doi: https://doi.org/10.1038/nrclinonc.2012.196
11. Lambin P, Zindler J, Vanneste B, van de Voorde L, Jacobs M, Eekers D, et al. Modern clinical research: How rapid learning health care and cohort multiple randomised clinical trials complement traditional evidence based medicine. *Acta Oncol* 2015; **54**: 1289–300. doi: https://doi.org/10.3109/0284186X.2015.1062136
12. Lambin P, Zindler J, Vanneste BG, De Voorde LV, Eekers D, Compter I, et al. Decision support systems for personalized and participative radiation oncology. *Adv Drug Deliv Rev* 2017; **109**: 131–53. doi: https://doi.org/10.1016/j.addr.2016.01.006
13. Wong AJ, Kanwar A, Mohamed AS, Fuller CD. Radiomics in head and neck cancer: from exploration to application. *Transl Cancer Res* 2016; **5**: 371–82. doi: https://doi.org/10.21037/tcr.2016.07.18
14. Chan MW, Yu E, Bartlett E, O'Sullivan B, Su J, Waldron J, et al. Morphologic and topographic radiologic features of human papillomavirus-related and -unrelated oropharyngeal carcinoma. *Head Neck* 2017; **39**: 1524 doi: https://doi.org/10.1002/hed.24764
15. Nesteruk M, Lang S, Veit-Haibach P, Studer G, Stieb S, Glatz S, et al. Tumor stage, tumor site and HPV dependent correlation of perfusion CT parameters and [18F]-FDG uptake in head and neck squamous cell carcinoma. *Radiother Oncol* 2015; **117**: 125–31. doi: https://doi.org/10.1016/j.radonc.2015.09.026
16. Buch K, Fujita A, Li B, Kawashima Y, Qureshi MM, Sakai O. Using texture analysis to determine human papillomavirus status of oropharyngeal squamous cell carcinomas on CT. *AJNR Am J Neuroradiol* 2015; **36**: 1343–8. doi: https://doi.org/10.3174/ajnr.A4285
17. Bogowicz M, Riesterer O, Ikenberg K, Stieb S, Moch H, Studer G, et al. Computed radiomics predicts HPV status and local tumor control after definitive radiochemotherapy in head and neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2017; **99**: 921–8. doi: https://doi.org/10.1016/j.ijrobp.2017.06.002
18. Shafiq-Ul-Hassan M, Zhang GG, Latifi K, Ullah G, Hunt DC, Balagurunathan Y, et al. Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels. *Med Phys* 2017; **44**: 1050–62. doi: https://doi.org/10.1002/mp.12123
19. Coroller TP, Grossmann P, Hou Y, Rios Velazquez E, Leijenaar RT, Hermann G, et al. CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiother Oncol* 2015; **114**: 345–50. doi: https://doi.org/10.1016/j.radonc.2015.02.015
20. Leijenaar RT, Nalbantov G, Carvalho S, van Elmpt WJ, Troost EG, Boellaard R, et al. The effect of SUV discretization in quantitative FDG-PET Radiomics: the need for standardized methodology in tumor texture analysis. *Sci Rep* 2015; **5**: 11075. doi: https://doi.org/10.1038/srep11075
21. Tibshirani R. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc Series B* 2011; **73**: 273–82. doi: https://doi.org/10.1111/j.1467-9868.2011.00771.x
22. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988; **44**: 837–45. doi: https://doi.org/10.2307/2531595
23. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiology* 2010; **21**: 128–38. doi: https://doi.org/10.1097/EDE.0b013e3181c30fb2
24. Rosenstein BS, West CM, Bentzen SM, Alsner J, Andreassen CN, Azria D, et al. Radiogenomics: radiobiology enters the era of big data and team science. *Int J Radiat Oncol Biol Phys* 2014; **89**: 709–13. doi: https://doi.org/10.1016/j.ijrobp.2014.03.009
25. Panth KM, Leijenaar RT, Carvalho S, Lieuwes NG, Yaromina A, Dubois L, et al. Is there a causal relationship between genetic changes and radiomics-based image features? An in vivo preclinical experiment with doxycycline inducible GADD34 tumor cells. *Radiother Oncol* 2015; **116**: 462–6. doi: https://doi.org/10.1016/j.radonc.2015.06.013
26. Larue RT, Defraene G, De Ruysscher D, Lambin P, van Elmpt W. Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *Br J Radiol* 2017; **90**: 20160665. doi: https://doi.org/10.1259/bjr.20160665
27. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. *Radiology* 2016; **278**: 563–77. doi: https://doi.org/10.1148/radiol.2015151169
28. Hatt M, Tixier F, Pierce L, Kinahan PE, Le Rest CC, Visvikis D. Characterization of PET/CT images using texture analysis: the past, the present… any future? *Eur J Nucl Med Mol Imaging* 2017; **44**: 151–65. doi: https://doi.org/10.1007/s00259-016-3427-0
29. Ajila V, Shetty H, Babu S, Shetty V, Hegde S. Human papilloma virus associated squamous cell carcinoma of the head and neck. *J Sex Transm Dis* 2015; **2015**: 1–5. doi: https://doi.org/10.1155/2015/791024
30. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. *Invest Radiol* 2015; **50**: 757–65. doi: https://doi.org/10.1097/RLI.0000000000000180
31. Zhao B, Tan Y, Tsai W-Y, Qi J, Xie C, Lu L, et al. Reproducibility of radiomics for deciphering tumor phenotype with imaging. *Sci Rep* 2016; **6**: 23428. doi: https://doi.org/10.1038/srep23428
32. Purohit BS, Ailianou A, Dulguerov N, Becker CD, Ratib O, Becker M. FDG-PET/CT pitfalls in oncological head and neck imaging. *Insights Imaging* 2014; **5**: 585–602. doi: https://doi.org/10.1007/s13244-014-0349-x
33. Leijenaar RTH, Carvalho S, Hoebers FJP, Aerts HJWL, van Elmpt WJC, Huang SH, et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma. *Acta Oncol* 2015; **54**: 1423–9. doi: https://doi.org/10.3109/0284186X.2015.1061214

34. Rietbergen MM, Snijders PJF, Beekzada D, Braakhuis BJM, Brink A, Heideman DAM, et al. Molecular characterization of p16-immunopositive but HPV DNA-negative oropharyngeal carcinomas. *Int J Cancer* 2014; **134**: 2366–72. doi: https://doi.org/10.1002/ijc.28580
35. Rios Velazquez E, Hoebers F, Aerts HJWL, Rietbergen MM, Brakenhoff RH, Leemans RC, et al. Externally validated HPV-based prognostic nomogram for oropharyngeal carcinoma patients yields more accurate predictions than TNM staging. *Radiother Oncol* 2014; **113**: 324–30. doi: https://doi.org/10.1016/j.radonc.2014.09.005
36. Fujita A, Buch K, Li B, Kawashima Y, Qureshi MM, Sakai O. Difference between HPV-positive and HPV-negative non-oropharyngeal head and neck cancer: texture analysis features on CT. *J Comput Assist Tomogr* 2016; **40**: 43–7.
37. Fischer CA, Kampmann M, Zlobec I, Green E, Tornillo L, Lugli A, et al. p16 expression in oropharyngeal cancer: its impact on staging and prognosis compared with the conventional clinical staging parameters. *Annals of Oncology* 2010; **21**: 1961–6. doi: https://doi.org/10.1093/annonc/mdq210
38. Cantrell SC, Peck BW, Li G, Wei Q, Sturgis EM, Ginsberg LE. Differences in imaging characteristics of HPV-positive and HPV-negative oropharyngeal cancers: a blinded matched-pair analysis. *AJNR Am J Neuroradiol* 2013; **34**: 2005–9. doi: https://doi.org/10.3174/ajnr.A3524